

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
12 June 2003 (12.06.2003)

PCT

(10) International Publication Number
WO 03/049035 A2

(51) International Patent Classification: G06T 7/40

(21) International Application Number: PCT/IB02/04819

(22) International Filing Date:
14 November 2002 (14.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/011,612 6 December 2001 (06.12.2001) US

(71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors: GUTTA, Srinivas, V., R.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). MIROSLAV, Trajkovic; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). COLMENAREZ, Antonio; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). PHILOMIN, Vasanth; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: GROENENDAAL, Antonius, W., M.; Internationale Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

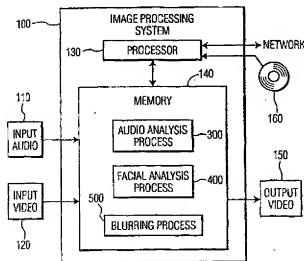
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR AUTOMATIC FACE BLURRING



(57) Abstract: An image processing system is disclosed that provides automatic face or skin blurring for images. All faces or skin can be blurred, or specific faces can be blurred. In one aspect of the invention, a particular face is blurred on an image or on a series of images in a video. Faces are determined in an image, and face matching is performed to match a particular face to faces in the image. If a match is found, the face or a portion of the face is blurred in the image. The blurring is performed on a portion of the image containing the particular face. Blurring may be performed through a variety of techniques. In another aspect of the invention, voice processing is used as an adjunct to or in place of face analysis to determine if a face in an image or series of images should be blurred. In another aspect of the invention, all faces or human skin in an image or series of images is blurred.

WO 03/049035 A2



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Method and apparatus for automatic face blurring

The present invention relates to image processing techniques, and more particularly, to a method and apparatus for automatically blurring faces and human skin in images.

5

Cameras are becoming ubiquitous in public and private areas. In public areas, cameras are being used for a variety of different and seemingly unrelated situations. For instance, retail stores use cameras to reduce shoplifting by customers and employees. As another example, many cities have cameras located at particular intersections or on particular highways, thereby providing real-time images of traffic. Additionally, another use for cameras in public places is face identification, which has been used at public events to capture wanted criminals.

Cameras are also becoming more popular in private areas, such as the home. Homemakers are using cameras to ensure that their children are properly cared for or for home security and surveillance purposes. Additionally, cameras are being integrated into home security systems, such that images from within a house can be transmitted to a monitoring facility if a home security system determines that something is amiss and that images should be sent.

The presence of cameras in these areas implicates privacy concerns. For example, in a shoplifting scenario, a criminal might be caught on video. However, other innocent shoppers may also be shown on the same video, and these people just happened to be in the store at the same time as the criminal. In a home situation, an intruder could enter the house when a security system is armed. The intruder, along with residents and guests in the house, could be caught in an image.

In these situations, there are people other than the criminal that are being shown on video. These people may or may not want to have their likeness on these videos.

There are some efforts to maintain privacy. For instance, on the television show "Cops," passersby and other individuals will have their faces blurred while the criminal and the police will not. This provides a way for bystanders to be blocked from view while the

criminal, and any others allowing their likeness to be used, is not. However, this blurring is manually added to images in a video. This is a relatively long and tedious task.

A need therefore exists for a method and apparatus for automatically modifying an image sequence to contain blurred facial images for those selected individuals, such as those people who do not wish to have their identity known.

Generally, an image processing system is disclosed that provides automatic face or skin blurring for images. All faces or skin can be blurred, or specific faces can be blurred.

In one aspect of the invention, a particular face is blurred on an image or on a series of images in a video. Faces are determined in an image, and face recognition is performed to match a particular face to faces in the image. If a match is found, the face or a portion of the face is blurred in the image. The blurring is performed on a portion of the image containing the particular face. Blurring may be performed through a variety of techniques.

In another aspect of the invention, voice processing is used as an adjunct to or in place of face analysis to determine if a face in an image or series of images should be blurred.

In another aspect of the invention, all faces or human skin in an image or series of images is blurred.

A more complete understanding of the present invention, as well as further features and advantages of the present invention, will be obtained by reference to the following detailed description and drawings.

FIG. 1 illustrates an exemplary image processing system in accordance with the present invention;

FIG. 2 illustrates a global view of the operations performed in accordance with the present invention;

FIG. 3 is a flow chart describing an exemplary implementation of the audio analysis process of FIG. 1;

FIG. 4 is a flow chart describing an exemplary implementation of the facial analysis process of FIG. 1; and

FIG. 5 is a flow chart describing an exemplary implementation of the blurring process of FIG. 1.

5 FIG. 1 illustrates an exemplary image processing system 100 in accordance with the present invention. According to one aspect of the present invention, the image processing system 100 interacts with input audio 110, input video 120, and Digital Versatile Disk (DVD) 160, and produces output video 150.

 The image processing system 100 may be embodied as any computing device,
10 such as a personal computer or workstation, containing a processor 130, such as a central processing unit (CPU), and memory 140, such as Random Access Memory (RAM) and Read-Only Memory (ROM). In an alternate embodiment, the image processing system 100 disclosed herein can be implemented as an application specific integrated circuit (ASIC), for example, as part of a video processing system. As shown in FIG. 1, and discussed further
15 below in conjunction with FIGS. 3 through 5, respectively, the memory 140 of the image processing system 100 includes an audio analysis process 300, a facial analysis process 400, and a blurring process 500. As known in the art, the processes 300, 400, and 500 can be combined into one process or made into even smaller processes. The processes 300, 400, and 500 are separated herein for ease of description and understanding.

20 Input audio 110 and input video 120 will usually be generated by a single camera with an attached microphone. However, the input audio 110 and input video 120 can be generated by separate devices. Additionally, input audio 110 is optional. Generally, input audio 110 and input video 120 will be digital. However, one or both of these signals can also be analog and image processing system 100 can use well known techniques to digitize the
25 signals.

 Generally, the audio analysis process 300 analyzes input audio 110 to determine if one or more particular speakers are speaking. If it is determined that one or more of the particular speakers are speaking, then all faces or potentially all human skin in an image or sequence of images is blurred. Additionally, a face, in an image, that corresponds to
30 the particular speaker can be blurred and tracked. Audio analysis also helps facial analysis because, if a particular speaker is found, the facial analysis then can start with the premise that a face for the particular speaker should be found in the image.

 The facial analysis process 400 determines human faces in images. The facial analysis process 400 determines locations in images for all human faces and can determine

particular faces. Optionally, the facial analysis process 400 can determine the presence of human skin in images. The blurring process 500 blurs regions of images. The blurring process 500 will essentially blur regions of images containing particular faces, all faces, or all skin.

- 5 As is known in the art, the methods and apparatus discussed herein may be distributed as an article of manufacture that itself comprises a computer readable medium having computer readable code means embodied thereon. The computer readable program code means is operable, in conjunction with a computer system, to carry out all or some of the steps to perform the methods or create the apparatuses discussed herein. The computer
- 10 readable medium may be a recordable medium (e.g., floppy disks, hard drives, compact disks such as DVD 160, or memory cards) or may be a transmission medium (c.g., a network comprising fiber-optics, the world-wide web, cables, or a wireless channel using time-division multiple access, code-division multiple access, or other radio-frequency channel). Any medium known or developed that can store information suitable for use with a computer
- 15 system may be used. The computer-readable code means is any mechanism for allowing a computer to read instructions and data, such as magnetic variations on a magnetic media or height variations on the surface of a compact disk, such as DVD 160.

- Memory 140 will configure the processor 130 to implement the methods, steps, and functions disclosed herein. The memory 140 could be distributed or local and the
- 20 processor 130 could be distributed or singular. The memory 140 could be implemented as an electrical, magnetic or optical memory, or any combination of these or other types of storage devices. The term "memory" should be construed broadly enough to encompass any information able to be read from or written to an address in the addressable space accessed by processor 130. With this definition, information on a network is still within memory 140 of
- 25 the image processing system 100 because the processor 130 can retrieve the information from the network.

- FIG. 2 illustrates an illustrative global view of the operations performed by the present invention. As shown in FIG. 2, audio analysis process 300 accepts and analyzes input
- 30 audio 110, using voice information 230. Audio analysis process 300 is described in detail in reference to FIG. 3. Voice information 230 allows audio analysis process 300 to determine if particular speakers are in input video 120. For instance, for a home security system, voices from inhabitants of the home will be collected for storage in the voice information 230 repository. For other systems, any person or persons who wish to have their privacy protected

through the face blurring techniques of the present invention can have their voice information stored in voice information 230.

When audio analysis process 300 determines that all or a portion of audio input 110 belong to one of the particular speakers, the audio analysis process 300 provides this information to facial analysis process 400. Facial analysis process 400 is described in more detail in reference to FIG. 4.

Facial analysis process 400 analyzes input video 120, which comprises a series of images 121, 122, and 123 in this example. Facial analysis process 400 examines each image 121, 122, and 123 to determine faces in the image and the locations and, generally, the approximate size of the faces. The facial analysis process 400 may also determine particular faces in the image. Optionally, the facial analysis process may simply examine images 121, 122, and 123 to search for human skin.

To determine specific faces in images 121, 122, and 123, facial analysis process 400 uses facial information 240. The facial analysis process 400 provides data to blurring process 500 to inform blurring process 500 what regions of an image to blur. Blurring process is described in greater detail in reference to FIG. 5. Blurring process 500 then blurs specific regions of each image 121, 122, and 123. In this manner, blurring process 500 produces output video 150, which comprises images 221, 222, and 223. Output video 150 comprises images that have regions blurred, if certain conditions are met.

In the example of FIG. 2, image 221 has region 220 blurred. Image 121 is a representation of an image before blurring takes place, and image 221 is a representation of this same image after blurring has been performed by blurring process 500. Region 220, in this example, blurs all of the face 210. However, portions, such as the eyes, of the face may be blurred instead. Additionally, all human skin in an image may be blurred, if desired.

FIG. 3 is a flow chart describing an exemplary implementation of the audio analysis process 300. As previously indicated, the audio analysis process 300 analyzes input audio 110, using voice information 230, to determine if specific speakers are in images. Generally, audio analysis process 300 performs speaker identification by using audio. There are a variety of techniques for speaker identification through audio. Each of the following U.S. Patents describes speaker identification by using audio, and each of these is hereby incorporated by reference: U.S. Patent No. 6,246,987, entitled "System for permitting access to a common resource in response to speaker identification and verification"; U.S. Patent No. 6,233,555, entitled "Method and Apparatus for Speaker Identification Using Mixture Discriminant Analysis to Develop Speaker Models"; U.S. Patent No. 6,141,644, entitled

"Speaker Verification and Speaker Identification Based on Eigenvoices"; U.S. Patent No. 6,029,124, entitled "Sequential, Nonparametric Speech Recognition and Speaker Identification"; and U.S. Patent No. 5,946,654, entitled "Speaker Identification Using Unsupervised Speech Models."

- 5 Method 310 starts in step 310, when the input audio 110 is analyzed. Generally, methods for speaker identification extract some type of information from the audio 110 (step 310), and this extracted information is compared (step 320) with stored voice information 230 from a speaker to determine a match within some predetermined accuracy.

- 10 For instance, in U.S. Patent 6,233,555, the voice information 230 is derived from two phases. In the enrollment phase a password utterance from a speaker is received multiple times. A transcription of the password utterance as a sequence of phones is obtained, and the phone string is stored in a database containing phone strings of other speakers in the group. In the training phase, the first set of feature vectors are extracted from each password utterance and the phone boundaries for each phone in the password transcription are obtained
- 15 using a speaker independent phone recognizer. A mixture model is developed for each phone of a given password from a speaker. Then, using the feature vectors from the password utterances of all of the speakers in the group, transformation parameters and transformed models are generated for each phone and speaker, using mixture discriminant analysis. The transformation parameters and transformed models are stored as voice information 230.

- 20 The process in U.S. Patent 6,233,555 then continues as follows. An utterance from a speaker is received from input audio 110 and a sequence of a first set of feature vectors are computed based on the received utterance (step 310). The first set of feature vectors are then transformed into a second set of feature vectors using transformations specific to a particular segmentation unit (step 310), and likelihood scores of the second set
- 25 of feature vectors are computed using speaker models (stored in voice information 230) trained using mixture discriminant analysis (step 310). The likelihood scores are then combined to determine an utterance score and the identity of a speaker is validated based on the utterance score (step 320).

- 30 Those skilled in the art will realize that there are a variety of techniques suitable for speaker identification.

 If an audio signature is found (step 330 = YES), which means that a speaker has been identified within a predetermined degree of accuracy, then information that indicates a particular speaker is output. This occurs in step 340.

It may be that it is possible to track a single speaker. For instance, if a camera is able to view an entire room where only one person is speaking, then the audio analysis could determine that the audio is from a speaker whose voice information has been stored in voice information 230. Once the identity of the speaker is determined, it is possible to use tracking techniques to track the person while the person remains in view of a camera.

It is also possible to determine that a person or persons in a room do not correspond to people who have their voice information stored in voice information 230. If another person enters the room whose audio signature can be determined in step 320, a system can then use tracking techniques to track this person. One commonly used tracking technique models the background, then determines what objects are moving by comparing new images with the modeled background. Such a technique is described in Elgammal A., Harwood D. and Davis L., "Non-Parametric Model for Background Subtraction," Proc. of the 6th European Conf. on Computer Vision, 751-767, Dublin, Ireland, June/July 2000, the disclosure of which is incorporated herein by reference.

In these situations, particular individual information 340 can include who the person is and what time he or she entered the view of a camera. This will help the facial analysis techniques, to be discussed below, because one or more people will be known to be present in the view of a camera.

Optionally, in step 340, the fact that a particular speaker was found is output instead of specific information about the person. This situation is particularly useful if all faces are to be blurred if the voice signature of one person is found. The latter can be a setting that a user can enable or disable.

If all of the voice information for all individuals have not been examined (step 350 = NO), then the audio analysis process 300 proceeds again in step 310. Optionally, step 310 is performed once each time audio analysis process 300 is performed, and step 350 begins again in step 320. If all the information for all individuals have not been examined (step 350 = NO), then the audio analysis process 300 stops.

As shown in FIG. 4, the facial analysis process 400 begins in step 405, when it is determined if a user has selected the option of blurring all faces in images. If this option has been selected (step 405 = YES), all faces are detected in step 410. The face detection may be performed in accordance with the teachings described in, for example, International Patent WO9932959, entitled "Method and System for Gesture Based Option Selection," assigned to the assignee of the present invention, Damian Lyons and Daniel Pelletier, "A Line-Scan Computer Vision Algorithm for Identifying Human Body Features," Gesture'99, 85-96

- France (1999), Ming-Hsuan Yang and Narendra Ahuja, "Detecting Human Faces in Color Images," Proc. of the 1998 IEEE Int'l Conf. on Image Processing (ICIP 98), Vol. 1, 127-130, (October, 1998); and I. Haritaoglu, D. Harwood, L. Davis, "Hydra: Multiple People Detection and Tracking Using Silhouettes," Computer Vision and Pattern Recognition, 5 Second Workshop of Video Surveillance (CVPR, 1999), each incorporated by reference herein.

- Many of these techniques require training, where the training is performed by having a camera view of an empty room to determine base data. Faces and movement are then more easily determined. These techniques are well known to those skilled in the art.
- 10 Detecting all faces 410 can be made easier by using previously stored locations of faces 415. Every time facial analysis process 400 can store current locations of faces in storage 415 and then use these locations when the facial analysis process 400 is executed the next time.

- Once all faces are found in step 410, the "locations" and areas of faces are estimated. The estimated locations are generally central locations of an outline of the face, and the outline of the face generally corresponds to the estimated area. The estimated locations can be influenced by user input 425. For example, a user could choose to have step 15 420 use ellipses to estimate areas of the face. The location of the face is then the center point of an ellipse and the ellipse can be tilted in such a way as to encompass the entire face.
- 20 Techniques for estimating head tilt and other parameters are given in the U.S. Patent application having serial number 09/898,139, entitled "Method and Apparatus for Interleaving a User Image in an Original Image Sequence," filed on July 3, 2001, by Gutta et al., the disclosure of which is incorporated by reference herein.

- As another example, a user may wish to have area estimated with rectangles or 25 other shapes. Rectangles may be useful, for instance, to ensure that the face and part of the hair are blurred in subsequent stages. As with ellipses, rectangles and other shapes can be appropriately tilted to provide the best estimate of facial area.

- After step 420, the facial analysis process 400 outputs blurring information in step 460. Such blurring information includes positions and areas of all faces, where the area 30 is generally described through mathematical shapes such as ellipses or rectangles.

If the option to blur all faces is disabled (step 405 = NO), then face information is retrieved from facial information repository 240. This occurs in step 430. In step 435, relevant faces are identified. A "relevant" face is simply a face that is to be blurred if the face is found to a predetermined degree of accuracy in an image. Generally, step 435

comprises two steps: face detection and then face recognition. For this version of step 435, generally movement is detected first, which involves background analysis, then do face detection on the object is performed. Another technique, which may be used in step 435, is to classify a object in the image and determine that the object is a human being. Classification of objects is described in Srinivas Gutta and Vasanth Philomin, "Classification of Objects through Model Ensembles," U.S. Serial No. 09/794443, filed February 27, 2001, the disclosure of which is incorporated by reference herein. After it is determined that the object is human, the system can detect a face and recognize the face.

As described above, there are a variety of techniques for face detection. There are also a variety of techniques for face recognition. The face recognition may be performed in accordance with the teachings described in, for example, Antonio Colmenarez and Thomas Huang, "Maximum Likelihood Face Detection," 2nd Int'l Conf. on Face and Gesture Recognition, 307-311, Killington, Vermont (October 14-16, 1996) or Srinivas Gutta et al., "Face and Gesture Recognition Using Hybrid Classifiers," 2d Int'l Conf. on Face and Gesture Recognition, 164-169, Killington, Vermont (October 14-16, 1996), incorporated by reference herein.

Audio analysis data 445, which is the output of audio analysis process 300, can also be used to determine that there is at least one person whose face should be blurred. This information can also contain information on who the person is. This enables step 435 to more easily perform face recognition. Additionally, if there is only one face in the view of the camera, then face recognition need not be performed. It should be noted that audio analysis data 445 may indicate that all faces are to be blurred once the audio signature of a single person is found. In this case, step 435 will act to find all faces and step 450 will act to provide location and area information for all faces.

It should also be noted that facial analysis process 400 can be simplified in certain situations by using previous locations of relevant faces 440. This will reduce the amount of processing by step 435 because the current location of a relevant face should be close to the previous location of a relevant face.

In step 450, the locations and areas of relevant faces are estimated and output information for relevant faces is output in step 460. Note that there may be no output information if no relevant faces are found. Method facial analysis process 400 then ends after step 460.

It should be noted that method 400 can optionally be used to determine human skin and then to blur the skin. This could be helpful, for instance, if a person in a video is

inappropriately dressed. An exemplary way of modifying method 400 to blur human skin is as follows. In step 405, it is determined that all human skin is to be blurred. In step 410, all human skin is found in an image. Human skin can be found through various techniques, such as the techniques described in Forsyth and Fleck, "Identifying Nude Pictures," Proc. of the
5 Third IEEE Workshop, Appl. of Computer Vision, 103-108, Dec. 2-4, 1996, the disclosure of which is incorporated by reference herein. Step 410 can use previously stored location 415.

In step 420, locations and areas of the human skin are determined, and this step may be affected by user preferences 425. The locations and areas are output in step 460. Again, it is possible to use shapes to describe the areas of the image containing skin.
10 Alternatively, it is also possible, in step 460, to create an output that will cause certain pixels in the image to be blurred. For example, a bitmap could be created that contains one bit for every pixel in the image could be created, where a "1" in the bitmap indicates that the pixel is to be blurred and a "0" in the image indicates that the pixel is not to be blurred. The present invention can use any technique suitable for indicating which pixels in an image should be
15 blurred.

FIG. 5 is a flow chart describing an exemplary implementation of the blurring process 500. As previously indicated, the blurring process 500 blurs regions of an image. Generally, these regions of the image correspond to faces. Blurring process 500 begins in
20 step 510, when blurring information is received. Such blurring information generally comprises sets of locations and areas. The locations are generally the center point of the area, and the area is generally described mathematically. This mathematics can include shapes and orientations for the shapes.

In step 520, the blurring process 500 blurs regions of an image by using the blurring information. The regions can be modified by user input. For instance, a user can
25 make the region be larger or smaller, or can choose different shapes. Previous blurring information is stored in previously blurred areas storage 525. This storage can be used to save user settings and to provide a "blur-only image" that can be combined with an input image to more easily provide the output blurred image. The blur-only image maps to the input image and provides locations to be blurred or perhaps pixel data to add to or subtract from the pixels
30 that already exist.

Additionally, the user input 425 can provide information on how the blurring should be performed. As is known in the art, there are a variety of techniques useful for blurring. For instance, an average could be determined from certain pixels. The average is then output to the pixels to be blurred.

It is to be understood that the embodiments and variations shown and described herein are merely illustrative of the principles of this invention and that various modifications may be implemented by those skilled in the art without departing from the scope and spirit of the invention.

CLAIMS:

1. A method comprising:
 - analyzing an image (121, 122, 123) using image processing techniques to find human skin (e.g., 210) in the image (121, 122, 123); and
 - blurring a region (e.g., 220) of the image (121, 122, 123) containing at least a
5 portion of the human skin (e.g., 210).
2. The method of claim 1, wherein the human skin (e.g., 210) is part of a human face (210).
- 10 3. The method of claim 1, wherein:
 - the step of analyzing further comprises analyzing an image (121, 122, 123) to find all human skin (e.g., 210) in the image (121, 122, 123); and
 - the step of blurring further comprises the step of blurring each region (e.g., 220) of the image (121, 122, 123) that contains human skin (e.g., 210).
- 15 4. A method comprising:
 - analyzing an image (121, 122, 123) to determine at least one face (210) in the image (121, 122, 123); and
 - blurring at least one region (e.g., 220) of the image (121, 122, 123)
20 containing at least a portion of the at least one face (210).
5. The method of claim 4, wherein:
 - the at least one face (210) comprises a plurality of faces;
 - the step of analyzing further comprises the step of identifying at least one
25 relevant face (210) from the plurality of faces; and
 - the step of blurring further comprises blurring a region (e.g., 220) of the image (121, 122, 123) containing at least a portion of the at least one relevant face (210).
6. The method of claim 4, wherein:

- the image (121, 122, 123) is one image (121, 122, 123) in a series of images from a video (120);

- the method further comprises the step of determining if sound in audio (110) corresponding to the video (120) can be assigned to a relevant person; and

5 - the step of blurring further comprises the step of blurring regions of each of the images when the audio (110) corresponds to a relevant person, each of the regions comprising one of the at least one faces, such that all faces in the images are blurred.

7. The method of claim 4, wherein:

10 - the image (121, 122, 123) is one image (121, 122, 123) in a series of images from a video (120);

- the method further comprises the step of determining if sound in audio (110) corresponding to the video (120) can be assigned to a relevant person; and

15 - the step of blurring further comprises the step of blurring a region (e.g., 220) of each of the images when the audio (110) corresponds to a relevant person, each of the regions comprising a face (210) that corresponds to the relevant person.

8. The method of claim 4, wherein a particular region (e.g., 220) comprises a face (210), and wherein the step of blurring further comprises blurring a portion of the region (e.g., 220) where the entire face (210) is blurred.

9. The method of claim 4, wherein the step of analyzing further comprises the step of determining, for each face (210) in the image (121, 122, 123), a location of the face (210).

25 10. The method of claim 9, wherein the step of determining a location further comprises the step of determining, for each location, an area.

11. The method of claim 10, wherein each area comprises a shape.

30

12. The method of claim 4, wherein:

the at least one face (210) comprises a plurality of faces;

the step of analyzing further comprises the step of identifying regions in the image (121, 122, 123) wherein the plurality of faces exist; and

the step of blurring further comprises blurring each of the regions so that each of the plurality of faces is blurred.

13. A system comprising:
5 a memory (140) that stores computer-readable code; and
a processor (130) operatively coupled to said memory (140), said processor configured to implement said computer-readable code, said computer-readable code configured to:
analyze an image (121, 122, 123) to determine at least one specified human
10 body feature in the image (121, 122, 123); and
blur at least one region (e.g., 220) of the image (121, 122, 123) containing at least a portion of the at least one specified human body feature.
14. The system of claim 13, wherein the at least one specified human body feature
15 comprises a portion of a face (210).
15. An article of manufacture comprising:
a computer-readable medium (e.g., 160) having computer readable code means embodied thereon, said computer-readable program code means comprising:
20 a step to analyze an image (121, 122, 123) to determine at least one specified human body feature in the image (121, 122, 123); and
a step to blur at least one region (e.g., 220) of the image (121, 122, 123) containing at least a portion of the at least one specified human body feature.

1/5

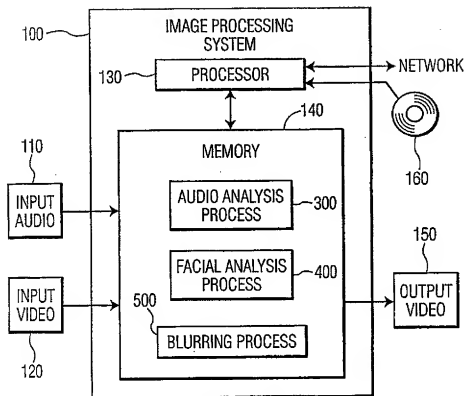


FIG. 1

2/5

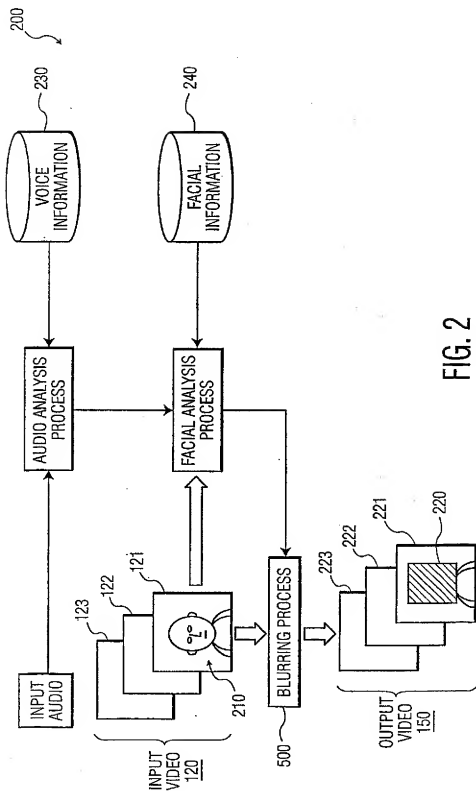


FIG. 2

3/5

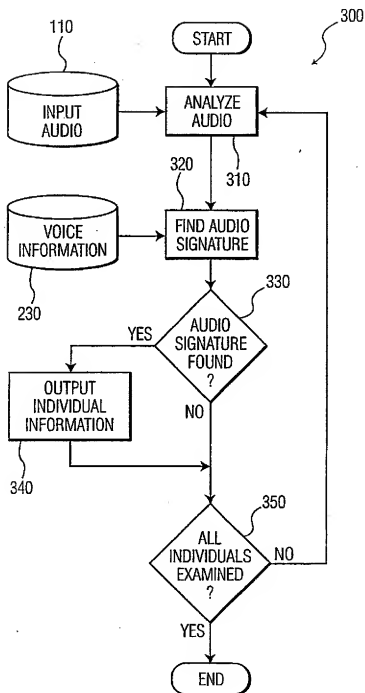


FIG. 3

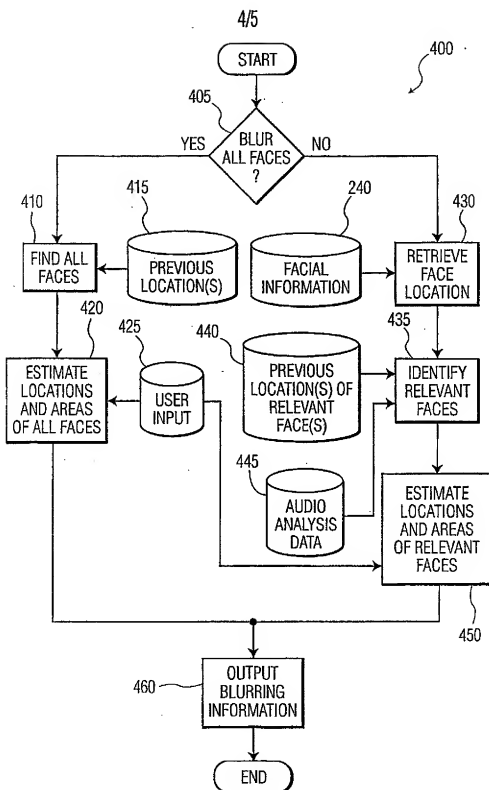


FIG. 4

5/5

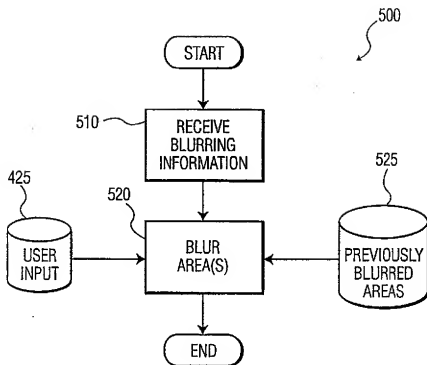


FIG. 5

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
12 June 2003 (12.06.2003)

PCT

(10) International Publication Number
WO 2003/049035 A3

- (51) International Patent Classification⁷: **G06K 9/00**
- (21) International Application Number:
PCT/IB2002/004819
- (22) International Filing Date:
14 November 2002 (14.11.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
10/011,612 6 December 2001 (06.12.2001) US
- (71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).
- (72) Inventors: GUTTA, Srinivas, V., R.; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). MIROSLAV, Trajkovic; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). COLMENAREZ, Antonio; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). PHILLOMIN, Vasanth; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (74) Agent: GROENENDAAL, Antonius, W., M.; Philips Intellectual Property & Standards, Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).
- (81) Designated States (national): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

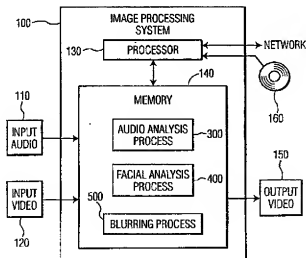
Published:

— with international search report

(88) Date of publication of the international search report:
17 June 2004

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR AUTOMATIC FACE BLURRING



(57) Abstract: An image processing system is disclosed that provides automatic face or skin blurring for images. All faces or skin can be blurred, or specific faces can be blurred. In one aspect of the invention, a particular face is blurred on an image or on a series of images in a video. Faces are determined in an image, and face matching is performed to match a particular face to faces in the image. If a match is found, the face or a portion of the face is blurred in the image. The blurring is performed on a portion of the image containing the particular face. Blurring may be performed through a variety of techniques. In another aspect of the invention, voice processing is used as an adjunct to or in place of face analysis to determine if a face in an image or series of images should be blurred. In another aspect of the invention, all faces or human skin in an image or series of images is blurred.

WO 2003/049035 A3



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 02/04819

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 G06K9/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the texts searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 1 107 166 A (MATSUSHITA ELECTRIC IND CO LTD) 13 June 2001 (2001-06-13)	4,9-15
Y	paragraphs '00031'-'00051', '00151', '00161', '00241', '00531', '00781', '00791'; figures 3A-3C, 9A-9C	1,2,5-8
A	US 5 524 162 A (LEVIEN RAPHAEL L) 4 June 1996 (1996-06-04) abstract column 1, line 54 - line 67 column 2, line 16 - line 25	1-15
Y	US 2001/000025 A1 (WOODFILL JOHN ET AL) 15 March 2001 (2001-03-15)	1,2,5,8
A	abstract; figures 2,4,5 paragraphs '00381'-'00431'	9-11
	-/-	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be of particular relevance

E earlier document but published on or after the international filing date

L document which may throw doubts on priority claims or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

I later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

A document member of the same patent family

Date of the actual completion of the international search

24 July 2003

Date of mailing of the international search report

01/08/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5816 Patentlaan 2
NL - 2200 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Kessler, C

INTERNATIONAL SEARCH REPORT

International Publication No.
PCT/IB 02/04819

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JAIN A ET AL: "A Multimodal Biometric System using Fingerprints, Face, and Speech" AUDIO- AND VIDEO-BASED BIOMETRIC PERSON AUTHENTICATION, XX, XX, 22 March 1999 (1999-03-22), pages 182-187, XP002211742 abstract; figure 1 sections 2.3 and 2.4	6,7
P, X	US 2002/081003 A1 (SOBOL ROBERT E) 27 June 2002 (2002-06-27) abstract; figure 4 paragraph '0028!	1,2,4, 13-15

INTERNATIONAL SEARCH REPORT

 International Publication No
 PCT/IB 02/04819

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 1107166	A	13-06-2001	EP 1107166 A2 JP 2001222719 A US 2001002932 A1	13-06-2001 17-08-2001 07-06-2001
US 5524162	A	04-06-1996	NONE	
US 2001000025	A1	15-03-2001	US 6188777 B1 AU 8584898 A EP 0998718 A1 WO 9906940 A1	13-02-2001 22-02-1999 10-05-2000 11-02-1999
US 2002081003	A1	27-06-2002	DE 10164201 A1 GB 2372168 A	25-07-2002 14-08-2002

